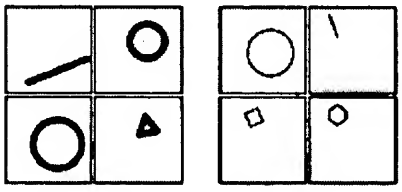
Can Hard AI Problems Foil Internet Interlopers?

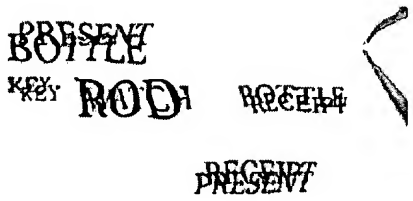
By Sara Robinson, SIAM News

In 1950, Alan Turing attempted a mathematical definition of the notion of a "thinking" machine and predicted that by 2002, scientists would have made significant progress in building such a machine. A computer can be said to think, Turing said, if it successfully answers questions in a procedure he called the "imitation game." The players in this game are two intelligent humans and the computer, placed in separate rooms and able to communicate only electronically. Human A gets to ask any questions he chooses of human B and the computer. If the computer manages to imitate a human so well that human A cannot tell which of the other players is human, the computer is said to be successfully " playing the imitation game." Turing went on to predict that within about fifty years, computers would be able to play this game so well that after five minutes of questioning "an average interrogator" would make the correct identification no more than 70% of the time.We know now that Turing was wrong---at least on the timing. While computers have exceeded expectations in most ways, we still don't know how to program a computer to imitate even a typical five-year-old. The field of artificial intelligence, which Turing's paper launched, has proved to be unexpectedly challenging.But one field's failure can be turned into another field's success. Just as the challenge of certain problems in number theory forms the basis for modern cryptography, some researchers are now making similar use of the hard aspects of AI. Hard AI problems have an advantage over hard number theory problems: Most human beings can solve them easily, but not so computers. This makes them ideal for the opposite of Turing's goal: the design of puzzles that create a barrier against computer programs designed to imitate humans---a problem that has a surprising number of applications on the Internet. Spearheading the effort to develop such puzzles is a group of researchers at Carnegie Mellon University, led by Manuel Blum, a cryptographer and theoretical computer scientist, and his PhD student Luis von Ahn. The group has dubbed its puzzles Captchas, an acronym for Completely Automated Public Turing Test to Tell Computers and Humans Apart.Among the Captchas is a visual character recognition puzzle called Gimpy. One instance of Gimpy is a picture containing seven distorted, overlapping words chosen at random from an 850-word dictionary. Solving the puzzle requires identifying three of the seven words and typing them into the box provided.Another Captcha puzzle, called Bongo, is based on a visual pattern recognition problem described in a 1951 book by M.M. Bongard. The user is presented with two sets of four patterned blocks, with the blocks in each set having some characteristic in common. The user is then asked to determine the set in which each of four additional blocks belongs. A third puzzle, with versions called Eco and Byan, requires the user to recognize sounds pronounced in varied tones in the presence of background noise.



Unlike computers, most humans can easily distinguish the distorted, overlapping words in Gimpy, one of several Captcha (Completely Automated Turing Test to Tell Computers and Humans Apart) puzzles designed by researchers at Carnegie Mellon University.

Bongo, another of the Carnegie Mellon researchers' Captcha puzzles, asks the user to assign each of four additional blocks to one of two four-block sets.



The Mathematical Capture of a Captcha

The Captcha project began in September 2000, when Yahoo! chief scientist Udi Manber told Carnegie Mellon computer science faculty about a "chat room" problem Yahoo! was experiencing. Companies were writing programs that could converse with humans in Yahoo! chat rooms, marketing the companies' products or gathering personal information.There were other problems as well. Yahoo! offers free services, such as e-mail accounts and Web-accessible data storage, recouping its expenses by selling advertising space. Spam companies, and other even less savory businesses, were causing headaches for Yahoo!, writing programs that could rapidly sign up for hundreds of free accounts that could then be used to send spam.Manber, who was formerly a professor of computer science at the University of Arizona, wanted a mechanism for screening out the automated interlopers, or 'bots, as such programs are called. Captchas were CMU's answer to Manber's problem.Unbeknown to the CMU group, in 1997 another group of researchers had come up with a similar solution to a problem with the AltaVista search engine. AltaVista enables users to enter Web pages missed by its Internet crawler program into its database. The ranking of a Web page is influenced by the number of other pages linking to it; sites were abusing the system by creating thousands of Web pages linked to their pages and writing programs to enter them into AltaVista's database.Seeking to ensure that only humans could enter pages into AltaVista's database, the group, from the Digital Equipment Systems Research Center, included an ad hoc character recognition test as part of the registration process. Andrei Broder, one of those researchers, credits the CMU team for defining the broader challenge and seeing its importance."Manuel [Blum] did a great thing by recognizing that this problem is much more than solving a nuisance for Yahoo! and AltaVista," Broder says. "It's useful for thinking about these issues in a completely new way."The CMU group decided that Captchas need to have certain basic properties: Most humans should be able to pass them, and current computer programs should not have a fixed probability of passing them, although a computer needs to be able to generate them at random from a very large database, and also to grade them. What's more, the database used to construct each instance of the puzzle should be public. Otherwise, says Blum, cracking the Captchas could be accomplished merely by breaking into the generating computer. And solving multiple instances of the puzzle gives some information about the database, he adds; why not make it open so the challenge is wholly in the puzzle?One of the Captchas, called Pix, initially made use of a large database of labeled images of objects---some photographs, others drawings or cartoons---to generate object recognition puzzles. Picking an object at random, along with six images of that object, Pix distorted the images and then asked the user to identify the common object and type the name into a box.Because the images were labeled in the database, an attacker could crack Pix just by gaining access to it, so Pix did not satisfy the criteria for a Captcha. To prevent such attacks, the images are now distorted by a random transformation after they have been selected from the database.Yahoo! recently integrated a simplified version of the word puzzle Gimpy, called E-Z Gimpy, into its registration process. The simplified version consists of a picture of only one word, distorted in one of several ways, and set into a noisy background.Use of the simplified version highlights another important property these puzzles must have if they are to be useful in practice: They must be extremely user-friendly. Manber was reluctant to deploy full Gimpy without further user testing, he says, fearing that it might deter a significant number of Yahoo!'s human registrants.Captchas should be out of reach even of current research efforts. A good Captcha is based on an AI problem so hard that researchers are confident that it won't be solved for a very long time.But what exactly is a hard AI problem?

Hard AI Defined

Certainly, a hard AI problem is related to a computationally hard problem. Given exponential time and space, many AI problems might become easy. But many computer scientists believe the human brain itself to be a sophisticated computing machine, and it's unlikely that any machine uses exponential-time algorithms.In fact, the CMU researchers don't consider the AI problems beneath the Captchas hard. Blum agrees with Turing's statement that all of them will eventually be conquered."Twenty years ago people said chess is hard, we'll never be able to write a program to play expert chess, and they did," Blum says. "I personally don't believe there is anything of an intellectual nature that we can do that computers cannot."Still, it would be useful to know that Captchas are as hard as the hard AI problems they're based on. Writing such theorems is common in cryptography, but in AI would require a precise definition of a hard AI problem, and it isn't clear how to create one. "We've decided not to follow that route," Blum says.Instead, in designing their Captchas, the researchers are using problems that AI researchers believe to be hard. If the Captchas are broken, it will be a loss in the battle against the 'bots, but an advance for the field of AI."We want people to break our Captchas because we want people to do good AI," says Luis von Ahn of Carnegie Mellon.Even if "good AI" is a long way off, incremental improvements in vision and sound recognition programs have many practical applications, says Nick Hopper, another member of the CMU Captcha group. Better vision programs could enable computers to scan in the entire Library of Congress, for instance, while sound recognition software with improved capabilities for filtering out background noise could enable researchers to design better hearing aids.

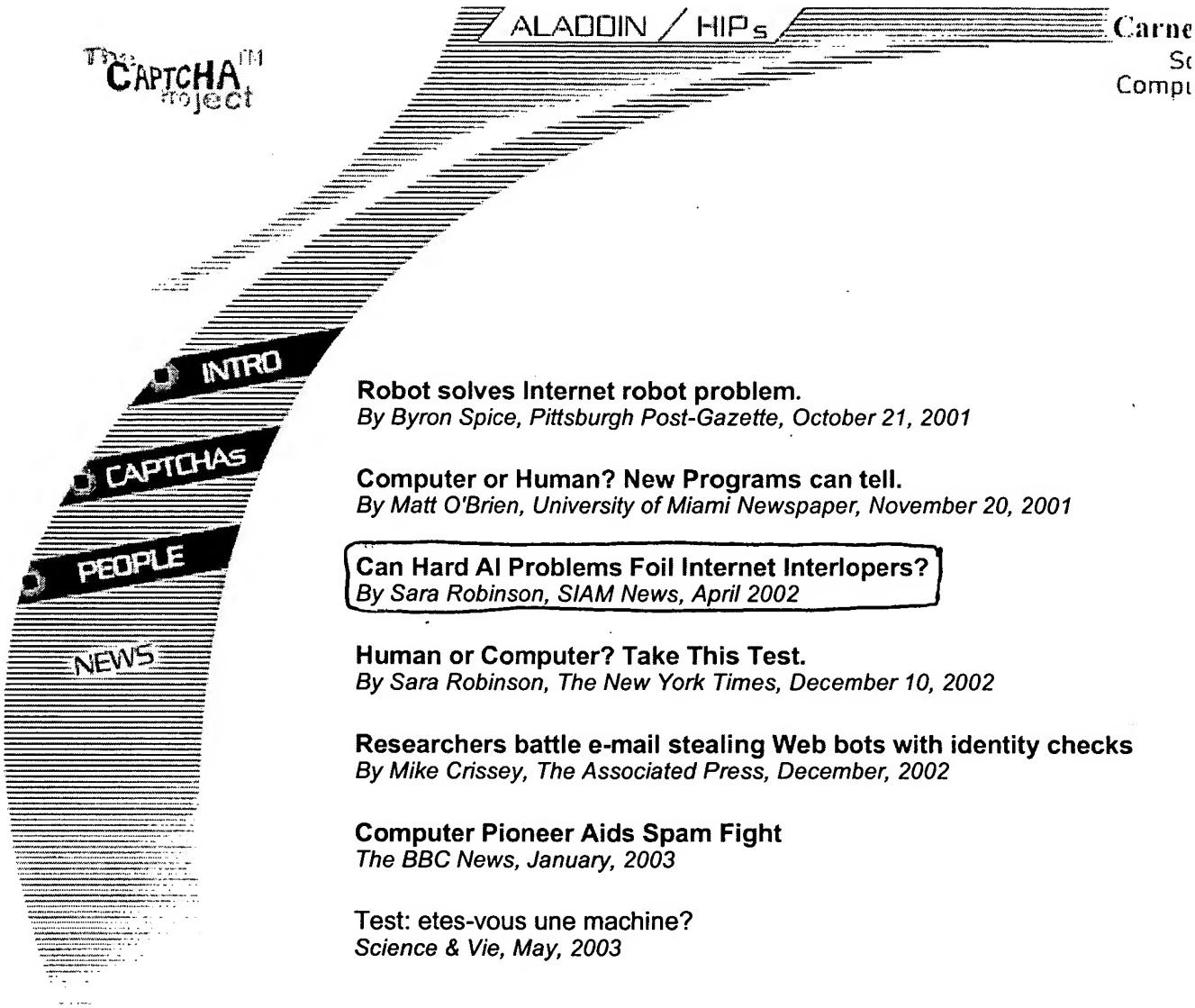
Why Captchas Are Hard

To test their programs, the Captcha researchers invited a number of prominent artificial intelligence researchers to examine the Captchas. Among them was Jitendra Malik, a computer vision specialist at the University of California, Berkeley.Malik judged AltaVista's version the most easily breakable and predicted that he could defeat E-Z Gimpy fairly easily as well. As for full Gimpy, "I don't know how to solve it, yet," he said.Malik believes that writing dedicated programs to solve each one of the puzzles is probably a feasible goal for a research project. The problem is that the time spent writing such a program would be wasted, because even a simple change to the puzzle would defeat the program.Indeed, cracks of early versions of several Captchas led, in most cases, to simple refinements that would prevent that particular attack. In an early version of one of the sound programs, for instance, words uttered in English were merely overlapped with words in other languages. Now, in the new version of Byan, the tones of the English words are distorted as well.Bongo, the visual pattern recognition puzzle, has been completely defeated, although only after years of effort. Harry Foundalis, a researcher at the University of Indiana, Bloomington, wrote a program that can quickly solve the underlying pattern recognition problem.Still, Henry Baird, a researcher at Palo Alto Research Center (PARC) and an expert in optical character recognition, warns that even writing dedicated programs for most of the Captchas is harder than it looks to those who haven't done it. The problems seem deceptively simple, he says, but getting such a program to work over the many possible variants is hard."If you try to build a machine vision system it has to be fully automatic, fast, accurate, and be able to handle a wide range of images," Baird says. "Usually, you can do two or three of these but in order to break a Captcha, you have to satisfy all at once."Baird has devised an OCR-inspired Captcha puzzle that incorporates many of the variations that are hard to resolve for current OCR programs. To create an instance of the puzzle, Baird's program picks a common English word and a typeface, then compresses or expands the text by a randomly chosen value. The program then adds salt-and-pepper noise to break up the boundaries and "thresholds" the images, setting a boundary value at which gray becomes black or white to break the characters into pieces.Empirical tests with UC Berkeley colleagues showed that three of the best existing commercial OCR systems are utterly baffled by this sort of text. Indeed, on most instances, they cannot guess a single letter, either correctly or incorrectly. Yet, using abilities whose mechanisms remain mysterious to computer scientists, most humans can read the text easily.

Do Captchas Do the Job?

As a way to articulate and make use of the hardness of Turing's imitation game, Captchas are clearly a success. But for Internet abuse problems, the value of Captchas is still unclear.Captchas have seemingly solved AltaVista's 'bot problem. The number of spurious database entries immediately decreased by 95% when the company included a Captcha in the registration process, Broder says.At Yahoo!, however, the account abuse problem has continued to increase. It's not clear whether the abusers are solving the Captchas or using other methods that get around them, Manber says. It seems that at least some of the companies attempting to open large numbers of Yahoo! accounts are using humans, who spend hours per day at the task, he adds."The real problem" for Yahoo!, Manber says, "is how to tell that somebody is the same person." Even if we are successful in designing Captchas that cannot be solved by computers, a person can sit for a whole day and solve hundreds of them. "We are trying other methods to identify that those solutions come from the same source."

Sara Robinson is a freelance writer based in Berkeley, California.



Robot solves Internet robot problem.
By Byron Spice, Pittsburgh Post-Gazette, October 21, 2001

Computer or Human? New Programs can tell.
By Matt O'Brien, University of Miami Newspaper, November 20, 2001

Can Hard AI Problems Foil Internet Interlopers?
By Sara Robinson, SIAM News, April 2002

Human or Computer? Take This Test.
By Sara Robinson, The New York Times, December 10, 2002

Researchers battle e-mail stealing Web bots with identity checks
By Mike Crissey, The Associated Press, December, 2002

Computer Pioneer Aids Spam Fight
The BBC News, January, 2003

Test: etes-vous une machine?
Science & Vie, May, 2003